

VIDEO PRODUCTION AND COMPACTION
WITH COLLAGE PICTURE FRAME USER INTERFACE

INVENTORS

Patrick Chiu
Shingo Uchihashi
John S. Boreczky
Jonathan T. Foote
Andreas Girgensohn
Lynn D. Wilcox

**CERTIFICATE OF MAILING BY "EXPRESS MAIL"
UNDER 37 C.F.R. § 1.10**

"Express Mail" mailing label number: **EL 622 697 664 US**
Date of Mailing: November 16, 2001

I hereby certify that this correspondence is being deposited with the United States Postal Service, utilizing the "Express Mail Post Office to Addressee" service addressed to **Box PATENT APPLICATION, Commissioner for Patents, Washington, D.C. 20231** and mailed on the above Date of Mailing with the above "Express Mail" mailing label number.


Austin Gloria

Signature Date: November 16, 2001

09992617.11601

ABSTRACT

[0082] A method, system, and apparatus for easily creating a video collage from a video is provided. By segmenting the video into a set number of video segments and providing an interface for a user to select images which represent the video segments and insert the selected images into a video collage template, a video collage may be easily created in a short amount of time. The system is designed to assign values to the video inserted in a video collage and compact the video based on these values thereby creating a small file which may be easily stored or transmitted.

5

09992617.11.1601

VIDEO PRODUCTION AND COMPACTION
WITH COLLAGE PICTURE FRAME USER INTERFACE

INVENTORS

Patrick Chiu
Shingo Uchihashi
John S. Boreczky
Jonathan T. Foote
Andreas Girgensohn
Lynn D. Wilcox

FIELD OF THE INVENTION

[0001] The present invention is related to the field of video production and editing.

BACKGROUND

[0002] Video cameras are becoming increasingly popular for home, office, and other types of use. For example, video cameras are frequently used to document family events such as vacations, weddings, and graduation ceremonies. In the office, video cameras are often used to record presentations, meetings, user studies and field work, and are often taken on business trips to record people, places, and activities.

[0003] Recently systems have been developed to interface personal computers with video cameras so that video can be stored in the memory of personal computers. As a result of this ability to transfer video from video tapes,

capture cards, and/or memory sticks, in a multitude of different formats to personal computers, there is an increasing demand for storing, viewing, editing, and transmitting video.

[0004] While unedited videos may often be viewed once or twice, they are often stored away since the interesting parts are generally intermixed with longer, less interesting regions. Further, unedited video is often of poor quality resulting from abrupt camera movement or too short or too long views, making it uninteresting to watch while waiting for the next interesting section.

[0005] In addition to unedited video often being long and boring, the amount of memory required to store the information is often rather large. A typical unedited video file may often range from a few bytes to hundreds of megabytes. Not only does unedited video often consume large amounts of memory, transmitting the video using e-mail or other electronic transmission systems is often difficult and slow.

[0006] Current systems for editing video, such as Adobe Premiere®, are designed for professional video producers and often take hours to assemble and mix video and audio tracks to produce an edited video. Other systems, such as the one described in commonly owned co-pending United States Patent Application Number 09/618,533 entitled "Interactive Custom Video Creation System," filed July 18, 2000, which is herein incorporated by reference, are easier to use but still require large amounts of user interaction to generate an edited video.

[0007] Until now, if a user wanted to electronically transmit a video, they either had to send the entire video or spend several hours creating an edited video

09992617 111601

and then send the edited video. While professionally editing video may be necessary in some circumstances, it is often undesirable and too time consuming.

[0008] Therefore, it is desirable to produce a system and method which segments video with a reduced amount of user input and provides the ability for a user to easily select portions of the video for viewing, storing, and or transmitting.

SUMMARY

[0009] Roughly described, the invention comprises a method for producing a video collage. The method may include segmenting a video into a plurality of video segments, providing a video collage template having at least one individual video frame, and associating a video segment from said plurality of video segments with said individual video frame of said video collage template. After the video segment has been associated with the video collage template the system generates a video collage from the video collage template and the associated video segment.

[0010] In one aspect, the invention comprises a video collage which includes a template having at least one individual video frame, and a representative image associated with a video segment, wherein said representative image is contained in said at least one individual video frame.

[0011] In a further aspect, the invention comprises a video collage user interface. The video collage user interface may include a video collage template having at least one individual video frame, a video segment template having a plurality of representative images, and a video segment selection device. Each of the representative images of the video segment template represent a video segment.

[0012] According to another aspect, the invention may comprise an apparatus for producing a video collage. The apparatus may include a processor and a processor readable storage medium in communication with the processor, containing processor readable program code. The processor readable program code
5 programs the apparatus to segment a video into a plurality of video segments, provide a video collage template having at least one individual video frame, and associate a video segment from said plurality of video segments with said individual video frame of said video collage template.

10 BRIEF DESCRIPTION OF THE DRAWINGS

[0013] The invention will be described with respect to the particular embodiments thereof. Other objects, features, and advantages of the invention will become apparent with reference to the specification and drawings in which:

[0014] Figure 1 illustrates an example of four video collage templates for
15 inserting and arranging video segments, according to an embodiment of the present invention;

[0015] Figure 2 illustrates a video segment template, according to an embodiment of the present invention;

[0016] Figure 3 illustrates an example of a user interface for creating a
20 video collage, according to an embodiment of the present invention;

[0017] Figure 4 illustrates an example of a video collage containing five selected representative images of video segments, according to an embodiment of the present invention;

[0018] Figure 5 is a generalized block diagram illustrating a computing device architecture suitable for implementing various embodiments of the present invention;

5 [0019] Figure 6 is a flow chart illustrating the steps for generating a video collage layout, according to an embodiment of the present invention;

[0020] Figure 7 is a flow chart illustrating the steps for generating video segments, according to an embodiment of the present invention;

[0021] Figure 8 is a flow chart illustrating the general steps for extracting and altering selected video segments; and,

10 [0022] Figure 9 is a table illustrating an example of video compaction, according to an embodiment of the present invention.

DETAILED DESCRIPTION

09992617 " 111601

[0023] Figure 1 illustrates an example of four video collage templates 100, 15 110, 120, and 130 for inserting and arranging video segments, according to an embodiment of the present invention. Each video collage template includes individual video frames for inserting video segments. For example, video collage template 100 contains individual video frames 101, 103, 105, and 107; video collage template 110 contains individual video frames 111, 113, and 115; video 20 collage template 120 contains individual video frames 121, 123, 125, 127, and 129; and video collage template 130 contains individual video frames 131, 132, 133, 134, 135, 136, 137, and 138. The individual video frames may be of any shape or size, constrained only by the size of the video collage template. For example, an

individual video frame may be oval, such as individual video frame 123, round, such as individual video frame 134, or any other shape. Additionally, there is no required number of individual video frames. A video collage user interface may have one, two, three, or any other number of individual video frames.

5 **[0024]** A video collage template may be provided by the system or may be created by a user. For example, a video collage template may be generated using Microsoft PowerPoint®, Microsoft Visio®, Microsoft Paint®, Corel Draw®, Deneba Systems, Inc.'s Canvas®, or any other type of drawing application, and stored in memory as an electronic file. Once a video collage template is created it
10 may be shared with other users and stored as an electronic file.

[0025] Figure 2 illustrates a video segment template 200, according to an embodiment of the present invention. Video segment template 200 includes representative images of associated segments of a video, such as representative images 201, 203, 205, 207, and 209. Video is segmented into a predetermined
15 number of segments and a representative image for each segment is selected and associated with that particular segment. The representative images may be displayed to a user using video segment template 200.

[0026] The representative image for each video segment may be determined in a variety of ways. For example, representative image 201 may be determined
20 by using the image contained in the first frame of the associated video segment. Alternatively, an image analysis algorithm may be applied to each frame of the video segment and a determination made as to which frame best represents the video segment and the image of the most representative frame displayed as the

representative image. In another example, a user may view the video segment and select a scene or frame from the video segment and have the image from the selected scene or frame used as the representative image.

[0027] Figure 3 illustrates a video collage user interface 300, according to an embodiment of the present invention. Video collage user interface 300 contains a user selected video collage template 310 and a video segment template 320 containing representative images of a segmented video, such as representative images 321, 323, 325, 327 and 329. Video collage user interface 300 may also include a group of controller mechanisms 301, 303, 305, 307, and 309 which may be used to perform various functions. For example, a user may select a representative image, such as representative image 323, from video segment template 320 using the arrow buttons 301, 303, 305, and 307 to position a selection box 350 around representative image 323 and activating a select mechanism 309. Additionally, a user may play a video segment associated with representative image 323 by again activating the select mechanism 309. Alternatively, once a representative image has been selected a user may position the selected representative image into a video frame of video collage template 310 using the controller mechanisms 301, 303, 305, 307, and 309.

[0028] Inserting a representative image into an individual frame of video collage template 310 will result in the representative image being displayed in the inserted video frame. The video segment associated with the representative image may be viewed by selecting the representative image. The associated video segment may be played in the individual video frame which contains the associated

representative image, may be played in an individual “window,” or may be played so as to fill an entire viewing area of a viewing device, such as display monitor 504 (Fig. 5).

[0029] The controller mechanisms 301, 303, 305, 307, and 309, and selection box 350, are used for illustrative purposes only and any other technique for selecting, viewing and positioning representative images and associated video segments may be used. For example, a user may perform these functions using any pointing device, such as a cursor control device 505 (Fig. 5), by selecting the representative image using a touch sensitive display screen, using keystrokes on a keyboard 506, etc.

[0030] Figure 4 illustrates an example of a video collage 400, according to an embodiment of the present invention. Video collage 400 is generated using video collage user interface 300 (Fig. 3). In generating video collage 400 a user selects and inserts representative image 321 into individual video frame 311, resulting in representative image 421 contained in video frame 411, selects and inserts representative image 323 into individual video frame 313, resulting in representative image 423 contained in video frame 413, selects and inserts representative image 325 into individual video frame 315, resulting in representative image 425 contained in video frame 415, selects and inserts representative image 327 into individual video frame 317, resulting in representative image 427 contained in video frame 417, and selects and inserts representative image 329 into individual video frame 319, resulting in representative image 429 contained in video frame 419, of video collage 400.

09992617 111601

[0031] Once video collage 400 has been generated it may be saved as an individual electronic file containing information relating to the arrangement and display of video collage 400 with selected representative images. Additionally, the file may contain video segments associated with the selected representative images.

5 [0032] In an alternative embodiment, the video collage or selected representative images may contain a link to a location where the associated video segments are stored. For example, if the video is available over the Internet, the electronic file may contain only information for displaying the video collage 400 with the representative image and a link to a location for each video segment
10 associated with a representative image. In such an example, if a representative image from video collage 400 is selected, the associated video segment is streamed over the Internet and displayed to the user. Including a link in the electronic file to each associated video segment provides the ability to store and transmit an electronic file of video collage 400 which is very small in size.

15 [0033] The electronic file of video collage 400 may be of any file type which is capable of being viewed, stored, and/or transmitted by a user. For example, an electronic file of video collage 400 may be saved as Moving Picture Experts Group ("MPEG") file, Audio Video Interleave ("AVI") file, Digital Video ("DV") file, QuickTime file, or any other video file format.

20 [0034] Once the video collage 400 has been generated, a user may display the video collage 400 on a web site, transmit the collage to others using email or other transmission devices, save the video collage as an electronic file, print a copy

of the video collage, create Digital Video Discs ("DVD"), Video Compact Discs ("VCD"), Compact Discs ("CD"), etc.

[0035] While the scope of the present invention extends far beyond use with the Internet, the Internet is used to exemplify embodiments of the present invention. Alternatively, embodiments of the present invention may be implemented on any type of end-to-end communication channel. For example, embodiments of the present invention may be implemented to function with an instant messaging system, an internal network, such as a Local Area Network ("LAN"), or over a peer-to-peer network.

10 [0036] Figure 5 illustrates a computing device architecture 500 suitable for implementing embodiments of the present invention. The computing device architecture 500 includes a processor 502, a storage device 503, Internet access equipment 510, such as a modem, and a display monitor 504. The architecture 500 may also include cursor control device 505, Random Access Memory ("RAM") 15 507, Read Only Memory ("ROM") 508, keyboard 506, and a graphics co-processor 509. All of the elements of the computing device architecture 500 may be tied together by a common bus 501 for transporting data between the various elements. The bus 501 typically includes data, address, and control signals.

[0037] Embodiments of the present invention are executable on any 20 computing device architecture such as the one 500 illustrated in Figure 5, but there is no limitation that this architecture is the only one which can execute embodiments of the present invention.

[0038] In an embodiment of the present invention, the storage device 503 may be an article of manufacture, such as a computer readable medium. For example, storage device 503 may be a magnetic hard disk, an optical disk, a floppy disk, CD-ROM (Compact Disk Read-Only Memory), RAM (Random Access Memory), ROM (Read-Only Memory), or other readable or writeable data storage technologies, singly or in combination.

[0039] Storage device 503 may include an operating system 511, such as Microsoft Windows®, Apple Macintosh OS®, or Unix®, wherein the operating system 511 is capable of executing programs or applications using computing device architecture 500. An embodiment of the present invention is implemented as video collage software program 512, and is stored on storage device 503.

[0040] As will be understood, embodiments of the present invention, such as video collage software program 512, may be in the form of a software program, a software object, a software function, a software subroutine, a software method, a software instance, a code fragment, a hardware operation or user operation, singly or in combination. Additionally, video collage software program 512 may be implemented using one, two, or any number of computing devices 500.

[0041] Figure 6 is a flow chart 600 illustrating a general view of the steps for generating a video collage layout, according to an embodiment of the present invention.

[0042] As one who is skilled in the art would appreciate, Figures 6, 7, and 8 illustrate logic boxes for performing specific functions. In alternative embodiments, more or fewer logic boxes may be used. In an embodiment of the

present invention, a logic box may represent a software program, a software object, a software function, a software subroutine, a software method, a software instance, a code fragment, a hardware operation or user operation, singly or in combination. For example, a logic box may represent a portion of video collage software program 512 illustrated in Figure 5, or any portion of other embodiments of the present invention.

[0043] In logic box 601 a user initiates the system and provides a video which they desire to use for generating a video collage layout. Upon receipt of a video, control is passed to logic box 603.

[0044] In logic box 603 the system generates a set of video segments from the provided video. The set of video segments will contain a predefined number of segments from the video. In an embodiment, a user may select how many video segments the video is to be divided into. Alternatively, a preset number may be used for segmenting the video. In still another embodiment, the system may determine the length of the provided video and select a segmentation number based on the length.

[0045] In logic box 604 a set of predefined video collage templates, such as video collage templates 100, 110, 120, and 130 (Fig. 1) are displayed to a user and one of the templates is selected. Alternatively, a user may generate and use their own video collage template.

[0046] In logic box 605 a video collage user interface is displayed to a user. The video collage user interface, such as the one 300 described with respect to Figure 3, includes the video collage temple selected in logic box 604 and a video

segment temple 200 (Fig. 2) containing representative images associated with each video segment generated in logic box 603.

[0047] A user, as illustrated by logic box 607, selects segments from the video segment template and inserts them into desired individual frames of the selected video collage template. Once each of the individual frames of the video collage template have been filled, or the user has indicated to the system that they have completed filling the individual frames of the video collage template, control is passed to logic box 609 and a video collage is generated from the selected video collage template and the selected representative images.

10 [0048] Figure 7 is a flow chart illustrating the steps for generating video segments illustrated as logic box 603 (Fig. 6), according to an embodiment of the present invention.

[0049] Video segmentation is begun in logic box 701 in response to providing a video. For example purposes we will describe video segmentation as illustrated by flow chart 700 using a video taken at an ice cream social which is originally fourteen minutes long.

[0050] Once a video is provided, initial video segmentation may be performed, as illustrated by logic box 703. Initial video segmentation may be performed using a number of well-known video segmentation techniques. For example, video may be segmented by "shots." A "shot" is defined as the time from which a video camera is turned to record to the time to which the record is turned off. Alternatively, algorithms for detecting scene changes may be used. Shots and scenes may also be segmented manually.

0999617.11604

[0051] Once the video is initially segmented, control is passed to logic box 704 and a determination is made as to whether the number of segments generated in logic box 703 is less than or equal to the number of segments ("N") which are to be provided to a user for selection. The number N of segments which are to be provided to a user for selection may be a predefined number, or a number provided by a user. If it is determined that the number of segments of video generated in logic box 703 is less than or equal to N, control is passed to logic box 707 and the video segments are used as the video segments to be provided to the user. However, if the number of video segments is not less than or equal to N, control is passed to logic box 705.

[0052] Referring to the ice cream social video example, the video is initially segmented into 179 segments using a typical color difference algorithm. Such an algorithm segments video based on color differences among frames. Any other typical segmentation algorithm may be used to initially segment the video. Since this is too many for a user to deal with quickly and easily, and N is set to 24, control is passed to logic box 705 where a smaller set of segments are generated.

[0053] In logic box 705, N video segments are selected from the set of video segments generated in logic box 703. For explanation purposes the set of N video segments may be referred to herein as $S(N)$. The set $S(N)$ will include the segments which are most representative of the entire video.

[0054] Video segment selection, as described in logic box 705, may be performed using a variety of video segmentation techniques. For example, N video segments may be selected using the Genetic Segmentation Algorithm ("GSA")

described in commonly owned co-pending United States Patent Application Number 09/611,389 entitled "A Genetic Segmentation Method for Data, Such As Image Data Streams," filed July 7, 2000, which is herein incorporated by reference.

Another technique for selecting segments is to look at histogram differences of adjacent frames and setting a threshold so that the places with differences exceeding the threshold are deemed to be segment boundaries.

[0055] Once N segments have been selected, control is passed to logic box 707 and the set of N video segments is used as the set of video segments provided to the user.

[0056] In an alternative embodiment, logic boxes 703 through 705 may be combined. In such an example, an initial segmentation will be performed on the video which returns a predetermined number N of video segments. Segmentation may be performed using GSA, where GSA is provided with N. GSA segments the entire video into N segments which are used as the video segments provided to the user in logic box 707.

[0057] Figure 8 is a flow chart illustrating the general steps for extracting and altering selected video segments for generating a video collage, as illustrated by logic boxes 607 and 609 (Fig. 6), according to an embodiment of the present invention.

[0058] The process begins at logic box 800 and control is passed to logic box 801, where a user selects representative images of associated video segments for each frame of a video collage template using any of the techniques described above. Once representative images for associated video have been selected and

inserted into individual video frames of a video collage template, control is passed to logic box 803.

[0059] In logic box 803 each video segment associated with the selected representative image is extracted from the video using any video extraction technique known in the art. For explanation purposes the set of extracted video frames may be referred to herein as $U(k)$, where $U(k) = \{s_0, s_1, s_2, \dots, s_{k-1}\}$, wherein k is the number of user selected segments. Once each video segment is extracted control is passed to logic box 805.

[0060] In logic box 805 an importance value is assigned to each extracted video segment. In an embodiment, the segments assigned to larger individual video frames in the video collage are assigned a higher importance value while video assigned to smaller video frames are assigned a lower importance value. Alternatively, a user may assign an importance value to the selected video segments.

[0061] In logic block 806 each segment is associated with a "feature vector" (also referred to herein as " $a(s)$ ") to characterize content activity. Examples of content activity features are: (1) length of the segment; (2) changes in a medium over time (visual, audio, and text), etc. Assigning a feature vector to each video segment provides another feature for adjusting video segments. Providing another feature removes the problem of linearly mapping the lengths of video segments to the sizes of the individual video frames. Linearly mapping video segments to frame size is a problem when a short segment is assigned to a large frame, forcing the other video segments to be much too short.

[0062] An example of defining a feature vector will be described with respect to visual changes between frames of a segment. Visual changes may be determined by finding color histogram differences using the formula $f(s)$ = number of different frames (F'') that appears in a segment s , where $f(s)$ is a value for the visual changes.

[0063] Motion caused by a pan will typically give a high $f(s)$, a zoom will generally yield a lower $f(s)$, a still camera on a still object will give a zero $f(s)$. Since a still camera on a moving object with a uniform background may generate a low $f(s)$, edge detection techniques may be performed on the images to properly adjust $f(s)$. Edge detection techniques are known in the art and will not be explained in further detail.

[0064] The feature vector may take into account both normalized visual changes and the length of the video segment by using the following formula to define the feature vector:

$$a(s) = (a_1, a_2) = \left(\frac{f(s)}{\max\{f(s) | s \in U(k)\}}, \frac{\text{length}(s)}{\max\{\text{length}(s) | s \in U(k)\}} \right)$$

and

$$|a(s)| = \sqrt{a_1^2 + a_2^2} \quad (1)$$

[0065] Feature vectors may be modified to have other components by adding other features. For example, audio features may be added by sampling the audio of a video segment (e.g. every 10 ms) and determining energy differences.

For text features, which may be beneficial to use when there is a transcript (e.g. close caption feed) or text extracted by Optical Character Recognition ("OCR") from presentation slides, the text stream may be sampled and a determination made as to the changes in the rate of text.

5 [0066] Once an importance value is assigned to each selected video segment (logic box 805) and a feature vector which takes into account the importance value and content activity is associated with the video segment (logic box 806) control is passed to logic box 807.

10 [0067] In logic box 807 the selected video segments may be altered to reduce the size of the resulting electronic file and/or to eliminate any remaining uninteresting portions of the video segments. To ease explanation of video alteration we will refer to the feature vectors as being normalized so that they are contained within a sphere having a radius bounded by the longest of the feature vectors.

15 [0068] Each feature vector is shortened so that the square of its length relative to the bounding radius is at most the size of its prescribed video frame relative to other video frames in the video collage template. Thus, by solving the following formula a determination may be made with respect to the video segments, associated feature vectors, and the corresponding individual video frame,
20 as to how much to compact each video segment with respect to the other selected video segments:

09992517 " 111601

$$\frac{|a(s)|}{|a_{\max}|} \leq \frac{A(s)}{A_{\max}}, \text{ or}$$

$$|a(s)| \leq |a_{\max}| \frac{A(s)}{A_{\max}} \quad (2).$$

[0069] Each video segment s that satisfies this constraint is left alone, and video segments which do not satisfy this constraint are adjusted until the constraint is true. In an embodiment, the video segments are adjusted by truncating the video from the end of the segment toward the beginning until the constraint is satisfied. Since $|a(s)|$ can be parameterized by time as $|a(s(t))|$, and is an increasing function of t , as t goes from 0 to T , where T is the length of the segment, there will be a smallest point t^* where the constraint is satisfied. One method for determining t^* is to recursively search the two halves of the term interval $[0, T]$.

[0070] The segment s assigned to the largest individual frame (with $A(s) = A_{\max}$) does not need to be adjusted (i.e. truncated), because this segment will always satisfy the constraint, i.e.:

$$|a(s)| \leq |a_{\max}| \frac{A(s)}{A_{\max}} = |a_{\max}|.$$

[0071] Hence, the resulting video collage will have the segment that a user prescribed as the most important (e.g. by placing the video segment in the largest individual frame) shown in its entirety. Video compaction techniques, other than

truncation, may also be used to reduce the size of the video segments. For example, another compaction technique is to time warp the segment (e.g. setting playback parameters so that it will speed up playback by a factor of p, where p is a number greater than 1.0).

5 [0072] Once the segments have been adjusted, as described above, control is passed to logic box 809 and a video collage is generated.

[0073] In an alternative embodiment, information retrieval ideas may be incorporated into the definition of the feature vector $a(s)$. Such information retrieval ideas may be term frequency, inverse document frequency, (" $tf*idf$ "). An example of this for video applications is described in *Summarizing video using a shot importance measure and a frame-packing algorithm*, Uchihashi, S., et. al., Proceedings ICAASP '99, vol. 6, pp. 3041-3044, which relies on hierarchical clustering of features, such as color; and in *A genetic algorithm for video segmentation and summarization*, Chiu, P., et al., Proceedings of ICME (IEEE Multimedia) 2000. In each of these papers, the value incorporated into the feature vector is referred to as "importance."

[0074] In such a variation we first define a set C_s to be those elements similar to a segment s in $U(k)$, where:

$$C_s = \{s \in U(k) \mid h(s, u) < \overline{dh} + \sigma, \forall u \in U(k)\} \quad \text{where } h(s, u) \text{ is defined by the}$$

20 histogram difference between the first frames of s and u in $U(k)$, $\overline{dh}(u)$ is defined by the histogram difference between the first frames of u and its next segment (in

time) in $U(k)$, and the mean \overline{dh} of dh and standard deviation σ of dh is computed over $U(k)$. In this formulation, we are only using visual features but it can be extended to include other types of features and media, such as audio. With the above description we define:

5
$$idf(s) = \log\left(\frac{1+k}{1+|C_s|}\right) .$$

[0075] $idf(s)$ can be normalized to be between 0 and 1 by dividing by a factor $C_{idf} = \log((1+k)/2)$. Next, this normalized equation is multiplied with the length term of each segment to obtain a variation of the feature vector:

$$a(s) = (a_1, a_2) = \left(\frac{f(s)}{\max\{f(s) | s \in U(k)\}}, \frac{length(s)}{L * C_{idf} * idf} \right)$$

10 [0076] In an alternative embodiment, to produce less truncation of the segmented video the constraint may be modified by determining the radius of each corresponding individual video frame. This can be done by taking the square root of the areas, or taking the square of the vector lengths for each video segment. In such an embodiment the following equation is utilized for determining segment
15 adjustment:

$$\frac{|a(s)|^2}{|a_{\max}|^2} \leq \frac{A(s)}{A_{\max}} . \quad (3)$$

[0077] Figure 9 illustrates an example of the amount of truncation which occurs using the above techniques for adjusting video, according to an embodiment of the present invention. In particular, the table in Figure 9 relates to the ice cream social video example. In this example, the system segments the ice cream video into twenty four segments and displays them through a video segment template included in the video collage user interface (Fig. 3). After user selection, the resulting video collage is illustrated as video collage 400 (Fig. 4). The segments 921, 923, 925, 927, and 929 correspond to the video segments associated with representative images 421, 423, 425, 427, and 429, respectively.

10 [0078] The table 900 illustrated in Figure 9 includes eight columns 901, 902, 903, 904, 905, 906, 907, and 908. Column 901 identifies each of the selected video segments 921, 923, 925, 927, and 929; Column 902 identifies the length in seconds of each selected video segment 921, 923, 925, 927, and 929; Column 903 identifies the area of each of the corresponding individual video frames 411, 413, 415, 417, and 419 (Fig. 4); Column 904 identifies the length of the computed feature vector for each selected video segment 921, 923, 925, 927, and 929; Column 905 identifies the percent each selected video segment 921, 923, 925, 927, and 929 is compacted using equation (2); Column 906 identifies the compacted length in seconds of each selected video segment 921, 923, 925, 927, and 929 using equation (2); Column 907 identifies the percent each selected video segment 921, 923, 925, 927, and 929 is compacted by using equation (3); and column 908 identifies the compacted length in seconds of each selected video segment 921, 923, 925, 927, and 929 using equation (3).

[0079] As can be seen from table 900, $a_{\max} = 1.18$ and relates to video segment 921, and $A_{\max} = 6.88$ and relates to individual video frame 415 which contains a representative image 425 of video segment 925. In this example, the five selected video segments consist of a total un-compacted length of 83.00 seconds but when compacted according to constraint equation (2) consist of a compacted length of 42.58 seconds, a compaction ratio of 49%. Referring to column 905, video segment 925 which corresponds to the largest individual video frame is not compacted because it is identified as having the highest importance. Video segment 921 which is associated with the largest feature vector however is in a smaller individual frame 411 and thus has a compaction ratio of 74% reducing the length of the segment from 28.00 seconds to only 7.28 seconds using constraint equation (2).

[0080] Alternatively, compacting the same five video segments using constraint equation (3) results in compacted length of 63.05 seconds. Referring to column 907, video segment 925 which corresponds to the largest individual video frame is not compacted because it is identified as having the highest importance. Video segment 921 which is associated with the largest feature vector, however, is in a smaller individual frame 411 (Figure 4) and thus has a compaction ratio of 45% reducing the length of the segment from 28.00 seconds to 15.30 seconds using constraint equation (3).

[0081] It should be understood that the particular embodiments described above are only illustrative of the principles of the present invention, and various modifications could be made by those skilled in the art without departing from the

scope and spirit of the invention. Thus, the scope of the present invention is limited only by the claims that follow.

0999647.11604